Narrative

Gilbert

12/4/2020

Brief substantive background / goal

"China Threat" discourse first appeared in the early 1990s in Japan, and then was borrowed by Western media to describe the growing challenger, China. As Sino-American competition gets increasingly intense, "China Threat" discourse also seems to get more prevalent in the Western media. According to power-transition theory and balance-of-threat theory, one would expect that as China grows stronger, threat from China perceived by U.S. will also increase, so discussion about "China Threat" should appear more and more. Is this really the case? This project aims to tackle this issue.

More specifically, I will webscrape articles about "China Threat" on media to reflect trend of "China Threat" discourse in the West, and then compare it with China's changing capabilities by running regression to see whether China's threat is associated with its increasing power.

To do this job, I need to use techniques including webscraping, regular expression, self-created function, if_else conditional statement, tiding data, transforming data, merging data, ggplot and statistical inference.

Collecting data

A natural candidate for webscraping is Western mainstream print media, like New York Times, or Washington Post, which can reflect the mainstream perception of "China Threat" in the West. Yet, since I am physically in mainland China, and Western media are blocked in China, I need to find a substitute, which becomes a big problem. Eventually, I decide to use People's Daily, the mouthpiece of Chinese Communist Party (CCP), for two reasons. First, given the aforementioned difficulty, this website is accessible in mainland China. Second, as the party's mouthpiece, CCP uses it as an important propaganda battle field to respond to Western's criticisms on it. Thus, the changing trend of articles about "China Threat" on People's Daily can reflect "China Threat" discourse in the West to some extent. Note: this way, my code will involve Chinese characters, a language that may be unfamiliaw to most classmates. I apologize for it beforehand. That said, after processing my data, the results I present are free of foreign language.

For China's changing capabilities, I use GDP and military expenditure data from World Bank database and PolityV index from Polity project to reflect China's economic, military capabilities, and authoritarianism, each corresponding to China as an economic, military, and political threat, respectively. A couple of words need to specify here. First, CINC score from Correlates of War dataset is usually used to reflect national capability, but CINC score only covers until 2012. Second, I admit that military expenditure is not a perfect proxy variable to reflect China's military capabilities, but China is a very closed country, whose military information is very difficult to get. Finally, it would certainly be better if I use relative power gap between China and U.S., instead of China's absolute power. For GDP, it works here, but the problem lies in military — America's military expenditure actually is getting increasingly larger than China. Yet, in reality, when observers claim China Threat, they actually focus exactly on China's growing capabilities alone, so absolute growth here can still do the job to some extent.

Cleaning / pre-processing data

First, I need to consider time scope. Since "China Threat" mainly emerged after the collapse of Soviet Union, I set the lower time limit on 1992. Since PolityV only covers data till 2018, I set the upper time limit here.

After scraping data from People's Daily, I can further categorize them into: China as a military threat, as an economic threat, and as a political threat. I detect words related to each of threat in each item respectively, and label the corresponding item as which kind of threat. (Articles include word "military" will be regarded as "China as a military threat"; articles include word "economy" and "development" will be regarded as "China as a military threat"; and articles include word "politics," "socialism," "ideology," and "political system" will be regarded as "China as a military threat"). Note that if one item includes both economic and military words, it will be regarded as both military and economic threat, so the total number of categarized threat doesn't equal to total number of articles. After doing it, since I want to plot it with respect to different types of threat, for convenience, I need to create a long format version. Finally, I get a dataframe with 108 rows and 3 columns:

In terms of GDP, military expenditure and PolityV data, I need to select data related to China first, because they are raw data including all countries from World Bank or Polity Project website. Since PolityV and "China Threat" data I created are in long format, while GDP and military expenditure are in wide format, before merging, I need to unify the formats first. With this done, I finally merge data about "China Threat" discourse, China's economic and military capabilities, and China's authoritarian level together.This returns me a dataframe with 27 rows and 8 columns:

Analysis and visualization

With the above data, I can first see how the number of articles on People's Daily that mentioned "China Threat" change over time, reflecting trend of "China Threat" discourse over time:

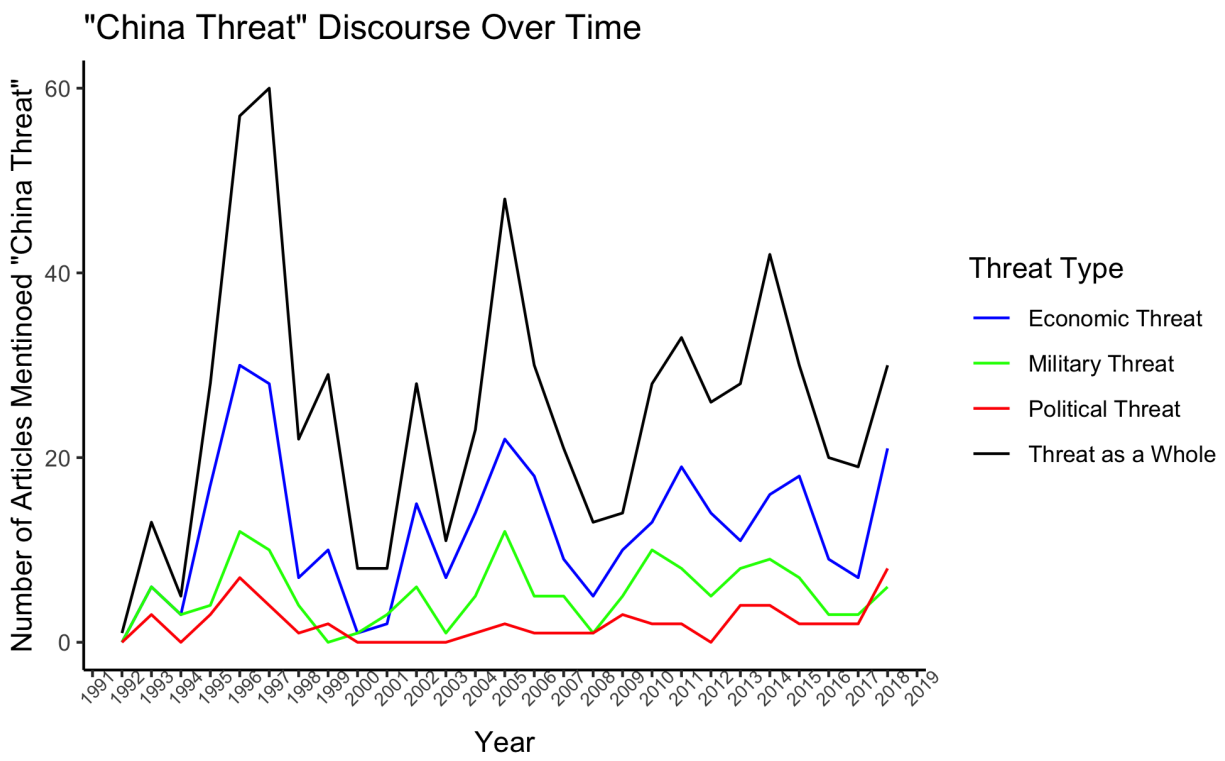

Figure 1: China Threat Over Time

It seems "China Threat" as a whole and other types of threat are related.

Then, let's see how China's economic and military power and its authoritarianism change over time:

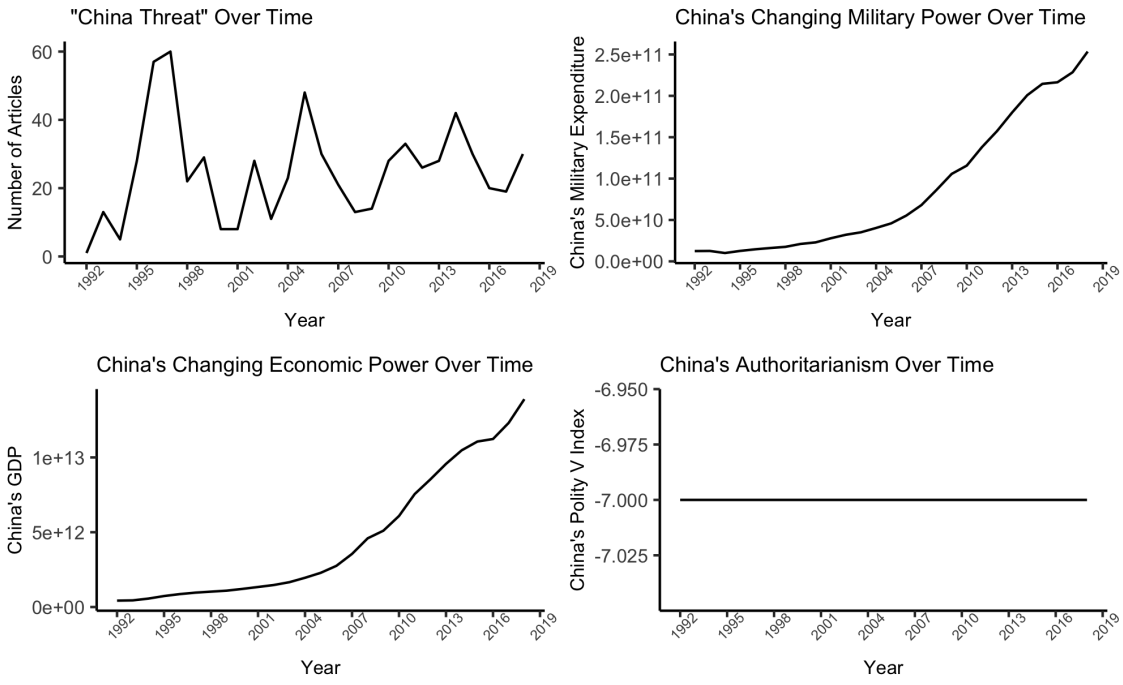

Figure 2: China Threat With Changing Capabilities

It seems "China Threat" and China's national capabilities not very correlated. Let's further do a plot to show whether China's specific type of threat is related to that kind of capability:

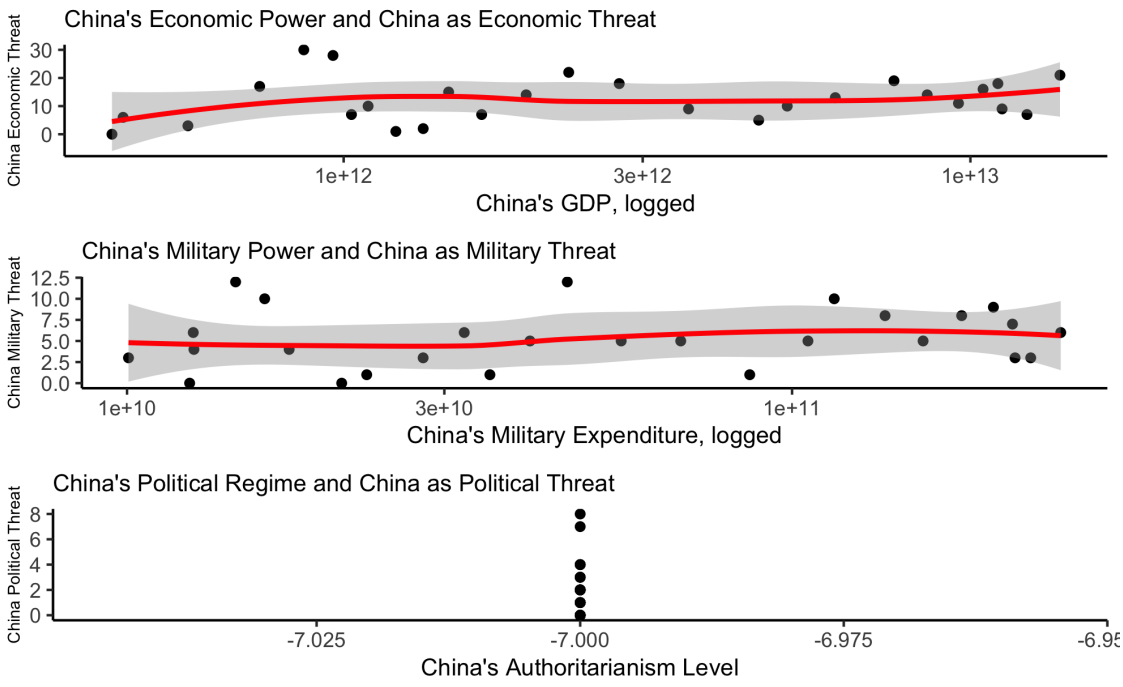

Figure 3: Regression Plot

Or do a statistical inference:

Regression Results

	Dependent variable:			
	Economic Threat (1)	Military Threat (2)	Political Threat (3)	Threat Discourse as a Whole (4)
China's GDP, logged	1.349			55.056**
	(1.353)			(20.394)
China's military expenditure, logged		0.602		-54.916**
		(0.613)		(21.030)
China's Authoritarianism				
Observations	27	27	27	27
R^2	0.038	0.037	0.000	0.243
Adjusted R^2	-0.0002	-0.001	0.000	0.180
Residual Std. Error	7.776 (df = 25)	3.417 (df = 25)	2.047 (df = 26)	13.184 (df = 24)
F Statistic	0.994 (df = 1; 25)	0.963 (df = 1; 25)		3.844** (df = 2; 24)

Note: *p<0.1; **p<0.05; ***p<0.01

Model 3 is blank because, as the above plot shows, regression fails, with Polity index being constant.

Model 4 is misleading, because "China Threat" discourse is driven by other factors (China's international action, I assume). However, since we only include economy, military, and polity here, regression does its job forcefully. Thus, here come the results we see.

Figure 4: Regression Table

This result is totally opposite to what we expected at first! Then, what's the problem? Go back to Steven Walt's balance-of-threat theory. Threat is made up of aggregate power, proximity, offensive power, and aggressive intention. Here, China's geography is fixed, and we have shown that China's growing power is not very correlated to "China Threat" discourse. This leaves us aggressive intention. By digging history, actually we can identity that the three highest peaks of "China Threat" exactly took place at times when China generated big international events.

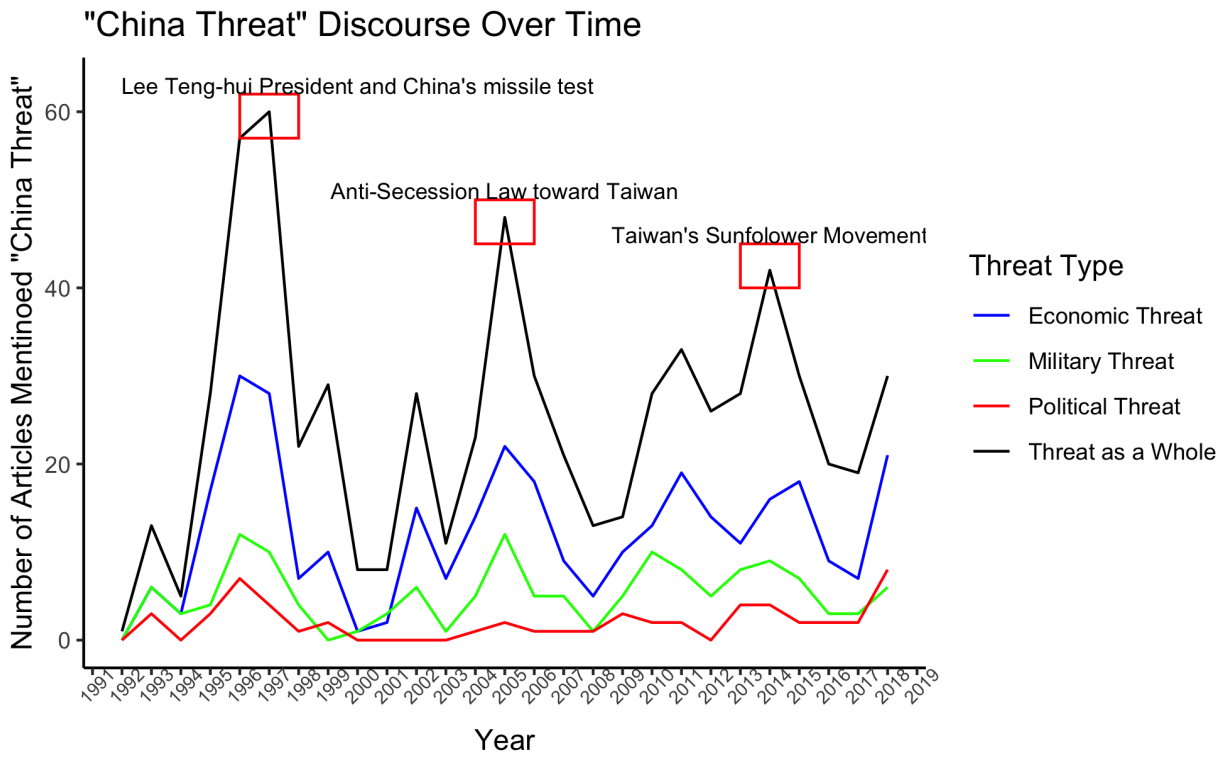

Figure 5: Plot Annotated

Thus, the result is not very departed from balance-of-threat theory, and we can conclude that "China Threat" discourse is not much about What China is, but What China does.

Future work

There are three main steps left for future studies:

First, the final conclusion being said, we can still see somewhat upward trend in the second half of the time period. Thus, it can also be that when China was weak, "China Threat" was more about what China did, but as China gets so powerful, "China Threat" becomes the result of combination of its growing power and its international behavior, thus, more consistent with Walt's theory. However, since a half of the time period specified here is too short, we cannot generally do a regression with so few observations. Thus, this waits for the future when we get more data.

Second, we can compare data from People's Daily with data from mainstream Western media to confirm whether "China Threat" discourse on People's Daily indeed reflects "China Threat" discourse in the West. If the answer is yes, then we can conclude that People's Daily, as a mouthpiece of Chinese Communist Party, does publish articles on "China Threat" as a response to "China Threat" discourse in the West. However, this work can only be done after I arrive in the states.

Finally, we can further test whether China's international behavior is associated with "China Threat" by doing such a trick: as long as China has an aggressive movement in that year, I add one point to a proxy variable. Thus, I can get a score to measure China's international behavior. Then, we can do a regression with "China Threat" and China's international behavior. If the result is significantly positive, then we can conclude that "China Threat" is indeed not a reflection of what China is, but what China does. This work requires detailed historical research, so it can leave for the futher work.